

251 AsnAlaGlyAsnLysAsnSerThrIleuLysLysLysAspSerProta 267  
 954 AAGCGCGGAAATTAATATCATATATTAACGTAAGAAAGAAATTCACCTGT 1003  
 267 IGIyGuileuTharSerLysThrAsnGlnAsnSerLysThrIleA 284  
 1004 AGGTGAATTTTAAACACGACGCAAAATATATCAAAATTTAAATATATA 1053  
 284 snTyfArGAspLeuThrIleGlyLysPheIleIleArgArgLysSer 300  
 1054 ATTATAGACATTATATATTTGAGAAATTTATATATAGAAAGAAAGTCA 1103  
 301 AsnSerGlnSerIleAsnAspIleValArgLysGlnAspThrIleY 317  
 1104 AATTCATCTATATATATATATATATATATATATATATATATATATA 1153  
 317 rLeuAspPheAsnLeuAsnGlnLysPheValArgValThrThrLys 334  
 1154 TCTAGATTTTAAATTAATCAAGAGTGAGAGATATATATATATATAT 1203  
 334 yPheLysLysGlnLysLysLysLysLysLysLysLysLysLysLys 350  
 1204 ATTTTAAAGAAAGAGAAAGAAATTTTGTCTTAACTATAGTATCT 1253  
 351 AspGluLeuThrAsnThrIleGlnIleLysGlnLysLysLysLys 367  
 1254 GATGAGTTTACATATCTATACAAATTAAGAAATGAGTAAAGACGCAAC 1303  
 367 rTyfSerCysGlnLeuLysLysLysLysLysLysLysLysLysLys 384  
 1304 ATATAGTGTCTAGTCTTTTAAAGAGATGAGAAAGTATGATGAGA 1353  
 384 IeGlyLeuIleGlyIleHisArgPheThrGluSerGlyIleValPheGlu 400  
 1354 TAGGATTTGATGATCTATCTCTTCTAGCAATCTGAAATTTATTTGA 1403  
 401 GluTyfLysAspTyfPheCysIleSerLysPheLysLysLysLys 417  
 1404 GAGTAAAGATTTATTTTGTATTAAGTAAATGCTTAAAGAGGTAA 1458  
 417 sArGlySerProTyfAsnLeuLysLysLysLysLysLysLysLys 434  
 1454 AAGGAACCATATATTTAAATTTGCAATGATGATGATGATGATGAT 1503  
 434 yAspGluGlyThrPheGlu 440  
 1504 AAGATGAAGGATGAGTGA 1523

seq\_name: /SISL1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:AAV30579

seq\_documentation\_block:  
 AAV30579 standard: DNA: 3876 BP.

AAV30579:  
 07-DEC-1998 (first entry)  
 Clostridium botulinum type B toxin gene from Danish strain.  
 Antitoxin: vaccine; neurotoxin; toxin B; intoxication; immunogen;  
 botulinum; BoTB; ds.  
 Clostridium botulinum serotype B Danish strain.  
 MO9808540-A1.  
 05-MAR-1998.  
 28-AUG-1997; 97MO-US15394.  
 28-AUG-1996; 96US-0704159.

(OPHI-) OPHIDIAN PHARM INC.  
 Thalley BS, Williams JA;  
 WPI: 1998-230234/20.  
 P-PSDB: AAW68392.  
 Host cell containing recombinant expression vector encoding  
 Clostridium botulinum type B or E toxin - useful to treat humans  
 and other animals at risk of intoxication with clostridial toxin.  
 Example 35; Page 291-296; 428pp; English.  
 This is the coding region of the Clostridium botulinum serotype B  
 (Danish strain) toxin gene that codes for a 1291-amino acid  
 polypeptide (see AAW68392). The C fragment (see AAW68394) of the  
 B toxin has been expressed as histidine-tagged protein in Escherichia  
 coli host cells. The invention relates to C. botulinum recombinant  
 toxin polypeptides. Methods are provided which allow for the  
 isolation of soluble recombinant proteins free of significant  
 endotoxin contamination. Preferred hosts for production of the  
 recombinant proteins are E. coli, insect cells and yeast cells.  
 The recombinant proteins are used as immunogens for the production  
 of vaccines and antitoxins that are useful in the treatment of  
 humans and animals at risk of intoxication with clostridial toxin.

Sequence 3876 BP; 1612 A; 370 C; 617 G; 1277 T; 0 other;

alignment\_scores:  
 Quality: 2333.00 Length: 440  
 Ratio: 5.326 Gaps: 0  
 Percent Similarity: 99.545 Percent Identity: 99.318

alignment\_block:  
 US-09-910-186A-8 x AAV30579

Align seg 1/1 to: AAV30579 from: 1 to: 3876

1 MetAlaAsnLysTyfAsnSerGluIleuAsnAsnIleIleuAsnLe 17  
 2554 AGCTTAATAATTAATTAATGCAAAATTTAAATATATATATCTTAATTT 2603  
 17 uArGlyTrLysAspAsnAsnLeuIleAspLeuSerGlyIleGlyAlaLys 34  
 2604 AAGATATAAGATATAATTAATTAATGATTTATCAGATATGCGCAAGC 2653  
 34 aGluValTyfAspGlyValGluLeuAsnAspLysAsnGlnPheLysLeu 50  
 2654 TAGAGGTATATGATGAGTCGAGCTTAATGATTAATCAATTTAAATTA 2703  
 51 ThrSerSerAlaAsnSerLysIleArgValThrGlnAsnGlnAsnIle 67  
 2704 ACTAGTTCAGCAAAATAGTAAGATTAGAGTGAAGTCAAAATCAAAATAT 2753  
 67 ePheAsnSerValPheLeuAspPheSerValSerPheThrIleArgIle 84  
 2754 ATTTATAGTGTGCTCTGATTTAGCGTTAGCTTTGATTAAGAAATAC 2803  
 84 rOLysTyfLysAsnAspGlyIleGlnAsnTyfIleHisAsnGluThr 100  
 2804 CTAATATTAAGAAATGATGATACAAATTTATATTCATTAATGAAATACA 2853  
 101 IleIleAsnCysMetLysAsnAsnSerGlyThrPheLysIleSerIleArg 117  
 2854 ATTAATTAATTTATGCAAAATTAATTCGGGCTGCAAAATATATATTAAGCG 2903  
 117 yAsnArgIleIleThrPheLeuIleAspIleAsnGlyLysThrLysSer 134  
 2904 TAATAGATATATAGGACTTAATGATTAATGAAATGAAACCAATGCG 2953  
 134 aPhePheGluTyfAsnIleArgGluAspIleSerGluTyfIleAsnArg 150

Tue Sep 3 14:28:26 2002

us-09-910-186

2954 TATTTTTTGAATATAACATAAGAGAAGATATATCAGAGTATATAAATAGA 3003  
151 TrpPhePheValThrIleThrAsnAsnLeuAsnAsnAlaLysIleTyrIl 167  
|||||  
3004 TGGTTTTTTGTAACATTACTAATAATTTGAATAACGCTAAAATTTATAT 3053  
167 eAsnGlyLysLeuGluSerAsnThrAspIleLysAspIleArgGluValI 184  
|||||  
3054 TAATGGTAAGCTAGAATCAAATACAGATATTAAGATATAAGAGAAGTTA 3103  
184 leAlaAsnGlyGluIleIlePheLysLeuAspGlyAspIleAspArgThr 200  
|||||  
3104 TTGCTAATGGTGAAATAATATTTAAATTAGATGGTGATATAGATAGAACA 3153  
201 GlnPheIleTrpMetLysTyrPheSerIlePheAsnThrGluLeuSerGl 217  
|||||  
3154 CAATTTATTGGATGAAATATTTTCAGTATTTTAAATACGGAATTAGTCA 3203  
217 nSerAsnIleGluGluArgTyrLysIleGlnSerTyrSerGluTyrLeuL 234  
|||||  
3204 ATCAAATATTGAAGAAAGATATAAAATTCAATCATATAGCGAATATTTAA 3253  
234 ysAspPheTrpGlyAsnProLeuMetTyrAsnLysGluTyrTyrMetPhe 250  
|||||  
3254 AAGATTTTGGGGAAATCCTTAAATGTACAATAAAGAATATTATATGTTT 3303  
251 AsnAlaGlyAsnLysAsnSerTyrIleLysLeuLysLysAspSerProVa 267  
|||||  
3304 AATGCGGGGAATAAAATTCATATATTAACATAAAGAAAGATTACCTGT 3353  
267 lGlyGluIleLeuThrArgSerLysTyrAsnGlnAsnSerLysTyrIleA 284  
|||||  
3354 AGGTGAAATTTTAACACGTAGCAAATATAATCAAATTTCTAAATATATA 3403  
284 snTyrArgAspLeuTyrIleGlyGluLysPheIleIleArgArgLysSer 300  
|||||  
3404 ATTATAGAGATTATATATTGGAGAAAAATTTATTATAAGAAGAAAGTCA 3453  
301 AsnSerGlnSerIleAsnAspAspIleValArgLysGluAspTyrIleTy 317  
|||||  
3454 AATTCTCAATCTATAAATGATGATATAGTTAGAAAAGAAGATTATATATA 3503  
317 rLeuAspPhePheAsnLeuAsnGlnGluTrpArgValTyrThrTyrLysT 334  
|||||  
3504 TCTAGATTTTAAATTTAAATCAAGAGTGGAGAGTATATACCTATAAAT 3553  
334 yrPheLysLysGluGluLysLeuPheLeuAlaProIleSerAspSer 350  
|||||  
3554 ATTTTAAGAAAGAGGAAGAAAAATGTTTTAGCTCCTATAAGTGATTCT 3603  
351 AspGluLeuTyrAsnThrIleGlnIleLysGluTyrAspGluGlnProTh 367  
|||||  
3604 GATGAGTTTACAATACTATACAAATAAAGAATATCATGAACAGCCAAC 3653  
367 rTyrSerCysGlnLeuLeuPheLysLysAspGluGluSerThrAspGluI 384  
|||||  
3654 ATATAGTTGTCAGTTGCTTTTAAAAAAGATGAAGAAAGTACTGATGAGA 3703  
384 leGlyLeuIleGlyIleHisArgPheTyrGluSerGlyIleValPheGlu 400  
|||||  
3704 TAGGATTGATTGGTATTTCATCGTTCTACGAATCTGGAATTGATTGTA 3753  
401 GluTyrLysAspTyrPheCysIleSerLysTrpTyrLeuLysGluValLy 417  
|||||  
3754 GAGTATAAAGATTATTTTGTATAAGTAAATGGTACTTAAAAAGAGTAAA 3803  
417 sArgLysProTyrAsnLeuLysLeuGlyCysAsnTrpGlnPheIleProL 434  
|||||  
3804 AAGGAACCATATAATTTAAATTTGGGATGTAATTGGCAGTTTATTCCTA 3853  
434 ysAspGluGlyTrpThrGlu 440  
|||||  
3854 AAGATGAAGGGTGGACTGAA 3873

SEQ ID 7

Page 1

186a-7.rge

1	742.8	55.4	4041	1	CLOBOTB	M81186 Clostridium
2	706.4	52.7	3876	1	AF295926	AF295926 Clostridi
3	706.4	52.7	3876	1	AF300465	AF300465 Clostridi
4	706.4	52.7	3876	1	AF300466	AF300466 Clostridi
5	706.4	52.7	11170	1	CBY13630	Y13630 Clostridium
6	696	51.9	3869	1	AF300467	AF300467 Clostridi
7	696	51.9	3869	1	AF300468	AF300468 Clostridi
8	696	51.9	3869	1	AF300469	AF300469 Clostridi
9	681.4	50.8	1326	1	CBO242628	AJ242628 Clostridi
10	663.4	49.5	4051	1	CBBONTB	X71343 C.botulinum
11	522.4	39.0	3509	6	A69701	A69701 Sequence 19
12	522.4	39.0	3509	6	BD009886	BD009886 Recombina
13	341	25.4	6862	1	CBNTNHTB	X87849 C.botulinum
14	263	19.6	1330	6	AR000029	AR000029 Sequence
15	263	19.6	1330	6	AR169140	AR169140 Sequence
16	263	19.6	1330	6	AX036243	AX036243 Sequence
17	263	19.6	1338	12	XXU22962	U22962 Synthetic b
18	263	19.6	1402	6	AR000030	AR000030 Sequence
19	263	19.6	1402	6	AR169141	AR169141 Sequence
20	263	19.6	1402	6	AX036246	AX036246 Sequence
21	242.6	18.1	1313	6	A58946	A58946 Sequence 6
22	241	18.0	1299	12	AF251281	AF251281 Synthetic
23	224.8	16.8	1084	1	CBBPOT	X70817 C.botulinum
24	220.2	16.4	3937	1	CBBONTG	X74162 C.botulinum
25	212.6	15.9	1084	1	CBBA129	X70814 C.botulinum
26	212.6	15.9	1084	1	CBBEKUND	X70819 C.botulinum
27	135.4	10.1	4067	1	CBNTOX	X73423 C.botulinum
28	134.2	10.0	3712	6	A49987	A49987 Sequence 4
29	131.4	9.8	1359	6	I28431	I28431 Sequence 3
30	129.4	9.6	4366	6	A42484	A42484 Sequence 12
31	129.2	9.6	3754	6	A37074	A37074 Sequence 17
32	129.2	9.6	3754	6	A42478	A42478 Sequence 6
33	129.2	9.6	3769	6	A37075	A37075 Sequence 18
34	129.2	9.6	3769	6	A49988	A49988 Sequence 5
35	129.2	9.6	4378	6	A42481	A42481 Sequence 9
36	123	9.2	3891	6	AR000031	AR000031 Sequence
37	123	9.2	3891	6	AR169142	AR169142 Sequence
38	123	9.2	3891	6	AX036248	AX036248 Sequence
39	123	9.2	4292	1	CBBOTAG	X52066 Clostridium
40	123	9.2	4835	1	CLONEUR	M30196 C.botulinum
41	103.2	7.7	1766	6	A49989	A49989 Sequence 6
42	96	7.2	1293	6	A58945	A58945 Sequence 5
43	96	7.2	4199	1	CLOBONT	L35496 Clostridium
44	96	7.2	4209	1	CBBONTFG	X81714 C.botulinum
45	92.6	6.9	3952	1	CBNTTD	X54254 Clostridium

## ALIGNMENTS

RESULT 1  
CLOBOTB 4041 bp DNA linear BCT 26-APR-1993  
LOCUS Clostridium botulinum neurotoxin type B (botB) gene, complete cds.  
DEFINITION M81186  
ACCESSION M81186.1 GI:144734  
VERSION botB gene; neurotoxin type B.  
KEYWORDS Clostridium botulinum DNA.  
SOURCE Clostridium botulinum  
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
Clostridium.  
REFERENCE 1 (bases 1 to 4041)  
AUTHORS Whelan, S.M., Elmore, M.J., Bodsworth, N.J., Brehm, J.K., Atkinson, T.  
and Minton, N.P.  
TITLE Complete nucleotide sequence of the Clostridium botulinum gene  
encoding the type B neurotoxin  
JOURNAL Unpublished (1991)  
FEATURES  
source Location/Qualifiers  
1..4041  
/organism="Clostridium botulinum"  
/db\_xref="taxon:1491"  
57..3932  
gene /gene="botB"

```
Query Match      55.4%   Score 742.8; DB 1; Length 4041;
Best Local Similarity    72.4%; Pred. No. 11e-170;
Matches 963; Conservative    0; Mismatches 367; Indels    0; Caps    0
```

Db 2610 ATGTTTAAATAATTAATAGCGAAATTTTAAATATATTTATCTTAAATTTAAAGATTAAG 2669  
 Oy 70 GACACAATCGATCGATCGTCTGCTGTGATCGTGTCTAAAGTGTGAAGTATCAAGCAAGTCTGATCAAC 69  
 Db 2670 GATATATATTTAATAGATTTTATCAGATATAGGGCGAAAGTAGAGTATATATGAGTCT 129  
 Oy 130 GAACGTGAAGACAAGAACCGATCTCAACCTGACCTCTCCGCTCAACTCTAAGATCCGTT 189  
 Db 2730 GAGCTTAAATGATAAAAATCAATTTAATTTAACTAGTTCAGCAAAATGTAAGATTGAGAGTG 2769  
 Oy 190 ACCGAGAATGAGACCTCATCTTCAACTCGATCTCTCGAGACTCTCTGTTCTCTCGG 249  
 Db 2790 ACTCAAAATCAGAAATATCATATTATTAATGCTGTTCTTGATTTTACGGTTAGCTTTTGG 2849  
 Oy 250 ATCTGATCCGGAATACAGAAGACGAGTGTCCGGAATCATCATCAACAATGAATACACC 309  
 Db 2850 ATTAGAAATACCTAAATATATAGAGTATGTTATCAAAAATTTATTTCAATATGAATATACA 2909  
 Oy 310 ATCTCAACTCGATGAAGAATAACTCTGTGTGAAGAGATCTCCATCCGCGCGTAACTCGTAA 369  
 Db 2910 ATTAATTAATTTGATCCAAAATAATTCGGGCGTGAATAATCTTATAGGGGTAAATAGGATA 2969  
 Oy 370 ATCTGACTCGATCGATATCAACGCTGAAGACCAAACTGTATCTCTCGATATCAACCAATC 429  
 Db 2970 ATATGCACTTTAATTTGATATTAATATGAAAAACCAATCCGCTATTTTGAATATACATA 3029  
 Oy 430 CGTGAAGACATCTGTGATCATCAATCAACGCTGTCTCTGTTACCAATCAACATACTGT 489  
 Db 3030 AGAAGAAATATATCAGAGTATTAATTAATAGTGGCTTTTGTAACTATTTACTATTAATTG 3089  
 Oy 90 AACCATCTAAATCTACCAACGCTGAACCTGGAATCTCAATACCGAATCAAGACT 549

Db	3090	AATAACCCCTAAATTTTATATTAATGTGAACCTACAAATCAAAATACAGATATTAAAGATATA	3149
QY	550	cgltgaagltacgcgtacgaacggtgaatcattccaaatccgaacgtgacgtgcacgcgtacc	609
Db	3150	AGAGAAGTATTCTCTAATGGTGAAATATATTTTAAATTTACATGGTGTATATGATAGACMA	3209
QY	610	cagltcatctcgtgataaataactctccatcttccaaacgcgaacgtctgcacgtccatcc	669
Db	3210	CAATTTATTGTGATGTAATAATTTACGTATTTTAAATACGAATTAAGTCAATCAATTAATTT	3269
QY	670	gaagaacggtacaaagatccagctctacccgcgaatcccgaaagacgttcggggtaaccg	729
Db	3270	GAAAGAAAGATTAATAATTCATATCAATATACCGCAATTTTAAAGATTTTGGGGAAATCCT	3329
QY	730	ctgagtctcaacaagaataactataatgttccaatgtcgtgtaacaaagaacgtctacatcaa	789
Db	3330	TTAATGTACAAATTAAGAATATTATATTGTTTAAATCGGGGAATTAATAATTCATATATTAA	3389
QY	790	ctgaagaagaagctctccggtctgtgtgaaatccctgcgtctccaaatcacacgaacgtct	849
Db	3390	CTAAGAAAGATTACCTCCTGATAGTGAAATTTTAAACACCTACCAAAATATATCAAAATTC	3449
QY	850	aaatcacatcaactacgcgcgcacgttacaatcggtgaagaagttcatalcatccgtccgaatct	909
Db	3450	AAATATATTAAATTTATAGAGATTTTATATATTGTGAGAAATAATTTATTAATGAAGAAAGCA	3509
QY	910	aactctagatccatcaatgatgatgcatactgcgtacgaagaagaactacatccctgcgtactc	969
Db	3510	AATTTCTCAATCTTAATAATGTATGATATGTAGAAAGAAAGATTAATATATCTAGATTTT	3569
QY	970	tccaacctgaatcaagaatgtgcgtgatacacaactcaagttactcgaagaagaagaagaa	1029
Db	3570	TTTAATTTTAATCAAGAGTGGAGAGATATTACTATAAATAATTTTAAAGAAAGAGAAAGAA	3629
QY	1030	aagcttctccgtgcgcgtacgtctctgtatcttcgaacgaactcacaagaacccatccgaatcaa	1089
Db	3630	AAATGTGTTTTAGCTCTCTTAAGTGAATCTGATGAATTTACATATCTATACAAATAAATAA	3689
QY	1090	gaatacgaacgaacgacgcgaacgtcctctgcgcacgtgcgtgttcaagaagaagaagaatct	1149
Db	3690	GAATTTGATGAACACGCCCAACATATATGTTGTCAGTTGCTTTTAAAAAAGATGACGAACGT	3749
QY	1150	actgaacgaatcgtgtcatcgtatcgtataccacgcgttctcgaacatcgtgtatcgtatcgaa	1209
Db	3750	ACTGATGATGATGAGATGATGATGATATTCATGTTCTTCTACGAATCTGGAAATCTATTTCGA	3809
QY	1210	gaatacgaagaagctcctctgcataccccaatggtacatcgaagaagaagttaacgcgaacgc	1269
Db	3810	GAGTATTAAGATTAATTTTGTATTAAGTAAATGATTAATAAAGAGGTAAAAAGAAACCA	3869
QY	1270	tacaacatcgaacggtgtgcgaatgtgcaatctcccgaaagaacgaagaagtttgaccgaa	1329
Db	3870	TATTAATTTTAATTTGGGATTTAATTTGGCACTTTATTTCTTAAGATGAAGGTGGACTGAA	3929
QY	1330	tagttaagaat 1339	
Db	3930	TATATTAACCT 3939	

RESULT	2
AF295926	
LOCUS	
DEFINITION	AF295926 3876 bp DNA linear BCT 02-SEP-2001
ACCESSION	Clostridium botulinum neurotoxin type B gene, complete cds.
VERSION	AF295926
KEYWORDS	AF295926.1 GI:15419707
SOURCE	.
ORGANISM	Clostridium botulinum. Clostridium botulinum Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium. 1 (bases 1 to 3876) Kirma,N., Ferreira,J.L. and Baumstark,B.R. Characterization of six type A strains of Clostridium botulinum
REFERENCE	
AUTHORS	
TITLE	

gb\_pat:AX088262 + 692.00 761.78 4.6e-34 4017 ! AX088262 Sequence 10 from  
 gb\_ba:AB037704 + 683.00 752.20 1.6e-33 3835 ! AB037704 Clostridium butyr  
 gb\_ba:AB037705 + 683.00 752.20 1.6e-33 3835 ! AB037705 Clostridium butyr  
 gb\_ba:AB037706 + 683.00 752.20 1.6e-33 3835 ! AB037706 Clostridium butyr  
 gb\_ba:AB037707 + 683.00 752.20 1.6e-33 3835 ! AB037707 Clostridium butyr

seq\_name: gb\_ba:CLOBOTB

seq\_documentation\_block:  
 LOCUS CLOBOTB 4041 bp DNA linear BCT 26-APR-1993  
 DEFINITION Clostridium botulinum neurotoxin type B (botB) gene, complete cds.  
 ACCESSION M81186  
 VERSION M81186.1 GI:144734  
 KEYWORDS botB gene; neurotoxin type B.  
 SOURCE Clostridium botulinum DNA.  
 ORGANISM Clostridium botulinum  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 Clostridium.

REFERENCE 1 (bases 1 to 4041)  
 AUTHORS Whelan, S.M., Elmore, M.J., Bodsworth, N.J., Brehm, J.K., Atkinson, T.  
 and Minton, N.P.

TITLE Complete nucleotide sequence of the Clostridium botulinum gene  
 encoding the type B neurotoxin

JOURNAL Unpublished (1991)

FEATURES Location/Qualifiers

source 1..4041

/organism="Clostridium botulinum"

/db\_xref="taxon:1491"

gene 57..3932

/gene="botB"

CDS 57..3932

/gene="botB"

/function="vertebrate neurotoxin"

/codon\_start=1

/transl\_table=11

/product="neurotoxin type B"

/protein\_id="AAA23211.1"

/db\_xref="GI:144735"

/translation="MPVTINNPNYNDPIDNNNIIMPEPPFARGTGRIYKAFKITDRIW  
 IIPERYTFGYKPEDFNKSSGIFNRDVEYDPPDLNTNDKKNIFLQTMIKLFNRKSK  
 PLGEKLLLEMIINGIPYLGDRVPLEEFNTNIASTVNKLISNPGEVERKKGIFANLII  
 FGPGPVLNENETIDIGIQNHFAFASREGFGGIMQKFCPEYVSFVFNQVQENKASIFNRR  
 GYFSDPALILMHLEIHLVHLGIGIKVDDLPIVNPNEKKFFMQSTDAIQAEELYTFGGQD  
 PSIIPTSDKSIYDKVLQNFGRGIVDRLNKVLVCISDPNININIKYKFKDKYKFVEDS  
 EGKYSIDVESFDKLYKSLMFGFTETNIAENYKIKTRASYFSOSLPPVKIKNLLDNEIY  
 TIEEGFNISDKDMEKEYRGONKAINKQAYEEISKEHLAVYKIQMKSKVAPGICIDVD  
 NEDLFFIADKNSFSDDLKNERIEYNTQSNYIENDFPINELILDITLISKIPLPSNT  
 ESILDFNVDVVPVYEQPAIKKIFTDENTIFQYLYSQTFFPLDIRDISLTIPLPSNT  
 NKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVNDFVIEANKSNTMDKIADISLIVPI  
 GLALNVGNETAKGNFENAFIAGASILLEFIPELLIPVVGAFLLSYIDNKKIITKI  
 DNALTKRNEKWSMDYGLIVAQLSTVNTQFYTIKEGMYKALNYQAQALEEIIKYRYNI  
 YSEKESNINIDFNDINSKLNENGINAIDNINNFINGCSVSYLMMKKMPLAVEKLLDF  
 DNTLKKNLLNYIDENKLYLIGSAEYKSKVKNYKLTIMPFDLSIYTNDTLIEMFNKY  
 NSEILNINIILNLRKDNLDLIDLSGYGAKVEYVYDGVELNDKNQFKLTSSANSKIRVTQ  
 QNIIFNSVFLDFSVSFIRIPKYKNDGIONYIHNEYTIINCMKNNSGWKISIRGNRII  
 WTLIDINGKTKSVFFEYNIREDISEYINRWFFVTITNNLNNAKIYINGKLESNTDIKD  
 IREVIANGELIFKLDGDIIDTQFIWMKYFSIFNTELSQSNIERYKIQSYEYLKDFW  
 GNPLMYNKEYYMFNAGNKNYSIKLKDSVPVGEILTRSKYNQNSKYINRYDLYIGKFI  
 IRRKSNSQSINDIVRKEDYIYLDFFNLNOEWRVYTYKFKKEEKLFLAPISDSDEF  
 YNTIQIKEYDEQPTYSQCLLFFKKDEESTDEIGLIGIHRFYESGIVFEYKDYFCISKW  
 YLKEVKKRPYNLKLGCNWFIPKDEGWTE"

BASE COUNT 1679 a 383 c 645 g 1334 t  
 ORIGIN

alignment\_scores:  
 Quality: 2340.00 Length: 440  
 Ratio: 5.330 Gaps: 0  
 Percent Similarity: 99.773 Percent Identity: 99.545

alignment\_block:  
 US-09-910-186A-8 x CLOBOTB

```

301 AsnserInserIleasnaspasPilevalrfgysgluspyrIleTy 317
3510 AATTCATCTATTAATGATGATATAGTTCGAAAGACGTTATATTA 3559
317 rLeuaspPhepheasLeuasnngIngluTrTPaYValYrThrTyLst 334
3560 TCTGATTTTTTTTAAATTAATCAAGAGCTGAGAGTATATACCTAAT 3609
334 YrPheLysLysGluGluLysLeuPheLeuAlaProIleSerAspSer 350
3610 ATTTTAAGAAAGAGAGAAATAATGTTTTTAACTCCTATTAAGATTC 3659
351 AspGluLeuTyraSnrThrIleGlnIleLysGluYrAspGluGlnPro 367
3660 GATGAGTTTACATACTATACAAATTAAGAAATATGATGACAGCCAC 3709
367 rTyRSetCysGlnLeuLeuPheLysLysAspGluGlnSerThrAspLui 384
3710 ATATAGTTGCAGTTCCTTTTTTAAAAAAGATGAAGAAAGTATGATAGA 3759
384 leGlyLeuIleGlyIleHisArgPheTyrcIuSerGlyIleValPheGlu 400
3760 TAGGATGATGTTGGATTCATCGTTTTCACAACTCGAATGTGATTTGAA 3809
401 GluTyrlLysaspyrPheCysIleSerLysTrpTyrlLeuLysGluValY 417
3810 GACTATTAAGATTAATTTTGTATAGATTAAGTAATGGTACTTAAGAGTAAA 3859
417 sArgLysProTyraSnrLeuLysLeuGlyCysAsnTrpGlnPheIlePro 434
3860 AAGGAACCATTAATTAATTTGGAGTGTAAATGGCACTTATATTCCTA 3909
434 ysAspGluGlyTrpThrGlu 440
3910 AAGATGAAGGCTGACTGAA 3929

seq_name: gb_ba:AF295926

seq_documentation_block:
LOCUS AF295926 3876 bp DNA linear BCT 02-SEP-2001
DEFINITION Clostridium botulinum neurotoxin type B gene, complete cds.
ACCESSION AF295926
VERSION AF295926.1 GI:15419707
KEYWORDS
SOURCE
ORGANISM
. Clostridium botulinum.
Clostridium botulinum
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridia
REFERENCE
1 (bases 1 to 3876)
Kirma,N., Ferreira,J.L. and Baumstark,B.R.
Characterization of six type A strains of Clostridium botulinum
that contain type B toxin gene sequences
Unpublished
2 (bases 1 to 3876)
Kirma,N., Ferreira,J.L. and Baumstark,B.R.
Direct Submission
Submitted (14-AUG-2000) Department of Biology, Georgia State
University, P.O. Box 4010, Atlanta, GA 30302-4010, USA
location/Qualifiers
1. 3876
/organism="Clostridium botulinum"
/isolate="1436"
/db_xref="taxon:1491"
/country="USA: Utah"
/note="Isolated from stool sample in 1977
type: AB"
1. 3876
/codon_start=1
/transl_table=11
/product="neurotoxin type B"
/protein_id="AAK97132.1"
/db_xref="GI:15419708"
FEATURES
source
CDS

```

251 AsnAlaGlyAsnLysAsnSerTyrIleLysLeuLysLysAspSerProva 267  
 954 AACCGGGAATTAATAATTCATATATTAACATAAGAAAGATTCACCTGT 1003  
 267 IGLYGLIleuLeuThrArgSerLysTyrAsnGlnAsnSerLysTyrIleA 284  
 1004 AGGTGAATTTTAAACAGTACCAATATATCAAAATCTAAATATATAA 1053  
 284 snTyrArgAsnLeuTyrIleGlyLysPheIleIleArgArgPsser 300  
 1054 ATTATAGCATTTTATATTTGGAGAAAAATTTATATTAAGAAAGATCA 1103  
 301 AsnSerGlnSerIleAsnAspPheValArgLysGlnAspTyrIleTyr 317  
 1104 AATCTCAATCTATATAATGATGATATAGTAAAGAAAGATTTATATA 1153  
 317 rLeuAspPhePheAsnLeuAsnGlnGluTyrArgValTyrThrTyrLys 334  
 1154 TCTAGATTTTATTTAATTAATCAAGAGTGAAGATATATACCTATAA 1203  
 334 rPheLysLysGluGluGluLysLeuPheLeuAlaProIleSerAspSer 350  
 1204 ATTTAAGAAAGAGAGAAAGAAATGTTTACCTCTATTAAGATTCAT 1253  
 351 AspGluLeuTyrAsnThrIleGlnIleLysGluLysTyrAspGluGlnPro 367  
 1254 GATGAGTTTACATACCTTTACCAATTAAGAAATGATGATGATGATG 1303  
 367 rTyrSerCysGlnLeuLeuPheLysLysAspGluGluSerThrAspGlu 384  
 1304 ATATAGTGTGCAATGCTTTTAAAAAGATGAAGAAAGATGATGATGA 1353  
 384 rLeuLeuIleGlyIleIleIleArgPheTyrGluSerGlyIleValPheGlu 400  
 1354 TAGGATTCATGATGATCATCGTTTTCACCAATCTGCAATGATTTGAA 1403  
 401 GluTyrLysAspTyrPheCysIleSerLysTyrTyrLeuLysGluVal 417  
 1404 GATATATTAAGATATTTTGTATTAAGTAATGATTAAGAAAGATGAA 1458  
 417 sArgLysProTyrAsnLeuLysLeuGlyCysAsnTyrGlnPheIlePro 434  
 1454 AAGGAAACCATTAATTAATTTGAGATGTAATGCGACTTATTCCTA 1503  
 434 yAspGluGluTyrTrpThrGlu 440  
 1504 AAGATCAAGGCTGACTGAA 1523

seq name: /SIDSI/gcdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:AAV30579

seq documentation\_block:

AAV30579 standard: DNA; 3876 BP.

AAV30579:

07-DEC-1998 (first entry)  
 Clostridium botulinum type B toxin gene from Danish strain.  
 Anti-toxin: vaccine; neurotoxin; toxin B; intoxication; immunogen;  
 Botulism; BoTB; ds.  
 Clostridium botulinum serotype B Danish strain.

W09808540-A1

05-MAR-1998.

28-AUG-1997; 97MO-US15394.  
 28-AUG-1996; 96US-0704159.

(OPHI-) OPHIDIAN PHARM INC.  
 Thalley BS, Williams JA.  
 MPI: 1998-230234/20.  
 P-PSDB; AAW68392.

Host cell containing recombinant expression vector encoding  
 Clostridium botulinum type B or E toxin - useful to treat humans  
 and other animals at risk of intoxication with clostridial toxin  
 Example 35; Page 291-296; 428pp; English.

This is the coding region of the Clostridium botulinum serotype B  
 (Danish strain) toxin gene that codes for a 1291-amino acid  
 polypeptide (see AAW68392). The C fragment (see AAW68394) of the  
 B toxin has been expressed as histidine-tagged protein in Escherichia  
 coli host cells. The invention relates to C. botulinum recombinant  
 toxin polypeptides. Methods are provided which allow for the  
 isolation of soluble recombinant proteins free of significant  
 endotoxin contamination. Preferred hosts for production of the  
 recombinant proteins are E. coli, insect cells and yeast cells.  
 The recombinant proteins are used as immunogens for the production  
 of vaccines and antitoxins that are useful in the treatment of  
 humans and animals at risk of intoxication with clostridial toxin.

Sequence 3876 BP; 1612 A; 370 C; 617 G; 1277 T; 0 other;

alignment\_scores:

Quality: 233.00

Length: 440

Ratio: 5.336

Gaps: 0

Percent Similarity: 99.545

Percent Identity: 99.318

alignment\_block:

US-09-910-186a-8 x AAV30579

Align seg 1/1 to: AAV30579 from: 1 to: 3876

1 MetAlaAsnLysTyrAsnSerGluIleLeuAsnIleIleLeuAsnIle 17  
 2554 ATGTTTAAATAAATAATATACGAAATTTTAAATAATATCTTAATTT 2603  
 17 uArgTyrLysAspAsnLeuLeuAsnSerGlyTyrGlyAlaLys 34  
 2604 AAGTATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2653  
 34 aGluValTyrAspGlyValGluLeuAsnAspLysAsnGlnPheLys 50  
 2654 TAGAGGATATGATGAGTGCATTAATTAATTAATTAATTAATTAAT 2703  
 51 ThrSerSerAlaAsnSerLysIleArgValThrGlnAsnGlnAsnIle 67  
 2704 ACTAGTTACCAATAATAGTATGATTAAGTACCAATAATGATTAAT 2753  
 67 ePheAsnSerValPheLeuAspPheSerValSerPheTyrIleArg 84  
 2754 ATTTAATAGTGTCTCTGATTTAGCGTTTACCTTTGGATTAAGAA 2803  
 84 rOlyTyrLysAsnAspGlyIleGlnAsnTyrIleHisAsnGluTyrThr 100  
 2804 CTAAATTAAGATGATGATTAATTAATTAATTAATTAATTAATTAAT 2853  
 101 IleIleAsnCysMetLysAsnAsnSerGlyTyrPheIleSerIleArg 117  
 2854 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2903  
 117 yAsnArgIleIleTyrPheLeuIleAsnIleAsnGlyLysTyrLys 134  
 2904 TATATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2953  
 134 aPhePheGluTyrAsnIleArgGluAspIleSerGluTyrIleAsnArg 150

Tue Sep 3 14:28:26 2002

us-09-910-186

2954 TATTTTTTGAATATAACATAAGAGAAGATATATCAGAGTATATAAATAGA 3003  
151 TrpPhePheValThrIleThrAsnAsnLeuAsnAsnAlaLysIleTyrIl 167  
|||||  
3004 TGGTTTTTTGTAAGTATTACTAATAATTTGAATAACGCTAAAAATTTATAT 3053  
167 eAsnGlyLysLeuGluSerAsnThrAspIleLysAspIleArgGluValI 184  
|||||  
3054 TAATGGTAAGCTAGAAATCAAATACAGATATTAAAGATATAAGAGAAGTTA 3103  
184 leAlaAsnGlyGluIleIlePheLysLeuAspGlyAspIleAspArgThr 200  
|||||  
3104 TTGCTAATGGTGAAATAATATTTAAATTAGATGGTGATATAGATAGAACA 3153  
201 GlnPheIleTrpMetLysTyrPheSerIlePheAsnThrGluLeuSerGl 217  
|||||  
3154 CAATTTATTTGGATGAAATATTTTCAATTTTAACTACGGAATTAAGTCA 3203  
217 nSerAsnIleGluGluArgTyrLysIleGlnSerTyrSerGluTyrLeuL 234  
|||||  
3204 ATCAAAATATTGAAGAAAGATATAAAATTCATCATATAGCGAATATTAA 3253  
234 ysAspPheTrpGlyAsnProLeuMetTyrAsnLysGluTyrTyrMetPhe 250  
|||||  
3254 AAGATTTTGGGGAATCCTTTAATGTACAATAAAGAAATATTATATGTTT 3303  
251 AsnAlaGlyAsnLysAsnSerTyrIleLysLeuLysLysAspSerProVa 267  
|||||  
3304 AATGCGGGAATAAAAAATTCATATATTAACTAAAGAAAGATTCACCTGT 3353  
267 lGlyGluIleLeuThrArgSerLysTyrAsnGlnAsnSerLysTyrIleA 284  
|||||  
3354 AGGTGAAATTTAACACGTAGCAATATAATCAAATCTAAATATATAA 3403  
284 snTyrArgAspLeuTyrIleGlyGluLysPheIleIleArgArgLysSer 300  
|||||  
3404 ATTATAGAGATTTATATATTGGAGAAAAATTTATTATAAGAGAAAGTCA 3453  
301 AsnSerGlnSerIleAsnAspAspIleValArgLysGluAspTyrIleTy 317  
|||||  
3454 AATTCCTCAATCTATAAATGATGATATAGTTAGAAAAGAGATTATATATA 3503  
317 rLeuAspPhePheAsnLeuAsnGlnGluTrpArgValTyrThrTyrLysT 334  
|||||  
3504 TCTAGATTTTTTAATTTAAATCAAGAGTGGAGACTATATACCTATAAAT 3553  
334 yrPheLysLysGluGluGluLysLeuPheLeuAlaProIleSerAspSer 350  
|||||  
3554 ATTTTAAGAAAGAGGAAGAAAAATGTTTTAGCTCCTATAAGTGATTCT 3603  
351 AspGluLeuTyrAsnThrIleGlnIleLysGluTyrAspGluGlnProTh 367  
|||||  
3604 GATGAGTTTACAATACTATACAAATAAAGAATATCATGAACAGCCAAC 3653  
367 rTyrSerCysGlnLeuLeuPheLysLysAspGluGluSerThrAspGluI 384  
|||||  
3654 ATATAGTTGTCAGTTGCTTTTTAAAAAGATGAAGAAAGTACTGATGAGA 3703  
384 leGlyLeuIleGlyIleHisArgPheTyrGluSerGlyIleValPheGlu 400  
|||||  
3704 TAGGATTGATTGGTATTCATCGTTTCTACGAATCTGGAATTGTATTGAA 3753  
401 GluTyrLysAspTyrPheCysIleSerLysTrpTyrLeuLysGluVally 417  
|||||  
3754 GAGTATAAAGATTATTTTGTATAAGTAAATGGTACTTAAAGAGGTAAA 3803  
417 sArgLysProTyrAsnLeuLysLeuGlyCysAsnTrpGlnPheIleProL 434  
|||||  
3804 AAGGAAACCATATAATTTAAATTTGGGATGTAATTGGCAGTTTATTCCTA 3853  
434 ysAspGluGlyTrpThrGlu 440  
|||||  
3854 AAGATGAAGGGTGGACTGAA 3873



186a-7.rge

1	742.8	55.4	4041	1	CLOBOTB
2	706.4	52.7	3876	1	AF295926
3	706.4	52.7	3876	1	AF300465
4	706.4	52.7	3876	1	AF300466
5	706.4	52.7	11170	1	CBY13630
6	696	51.9	3869	1	AF300467
7	696	51.9	3869	1	AF300468
8	696	51.9	3869	1	AF300469
9	681.1	50.8	1326	1	CBO242628
10	663.4	49.5	4051	1	CBONTB
11	522.4	39.0	3509	6	A69701
12	522.4	39.0	3509	6	BD009886
13	341	21.4	6862	1	CBNTNHTB
14	263	19.6	1330	6	AR000029
15	263	19.6	1330	6	AR169140
16	263	19.6	1330	6	AX036243
17	263	19.6	1338	12	XXU22962
18	263	19.6	1402	6	AR000030
19	263	19.6	1402	6	AR169141
20	263	19.6	1402	6	AX036246
21	242.6	18.1	1313	6	A58946
22	241	18.0	1299	12	AF251281
23	224.8	16.8	1084	1	CBPPTG
24	220.2	16.4	3937	1	CBPNTG
25	212.6	15.9	1084	1	CBP129
26	212.6	15.9	1084	1	CBP129
27	135.4	10.1	4067	1	CBNTXAX
28	134.2	10.0	3712	6	A49987
29	131.4	9.8	1359	6	I28431
30	129.4	9.6	4866	6	A42484
31	129.2	9.6	3754	6	A37074
32	129.2	9.6	3754	6	A42478
33	129.2	9.6	3769	6	A37075
34	129.2	9.6	3769	6	A49988
35	129.2	9.6	4378	6	A42481
36	123	9.2	3891	6	AR000031
37	123	9.2	3891	6	AR169142
38	123	9.2	3891	6	AX036248
39	123	9.2	4292	1	CBBOTAG
40	123	9.2	4835	1	CLONEUR
41	103.2	7.7	1766	6	A49989
42	96	7.2	1293	6	A58945
43	96	7.2	4199	1	CLOBONT
44	96	7.2	4209	1	CBONTFG
45	92.6	6.9	3952	1	CBNTTD

M81186 Clostridium  
AF295926 Clostridi  
AF300465 Clostridi  
AF300466 Clostridi  
Y13630 Clostridium  
AF300467 Clostridi  
AF300468 Clostridi  
AF300469 Clostridi  
AJ242628 Clostridi  
X71343 C.botulinum  
A69701 Sequence 19  
BD009886 Recombina  
X87849 C.botulinum  
AR000029 Sequence  
AR169140 Sequence  
AX036243 Sequence  
U22962 Synthetic b  
AR000030 Sequence  
AR169141 Sequence  
AX036246 Sequence  
A58946 Sequence 6  
AF251281 Synthetic  
X70817 C.botulinum  
X74162 C.botulinum  
X70814 C.botulinum  
X70819 C.botulinum  
X73423 C.botulinum  
A49987 Sequence 4  
I28431 Sequence 3  
A42484 Sequence 12  
A37074 Sequence 17  
A42478 Sequence 6  
A37075 Sequence 18  
A49988 Sequence 5  
A42481 Sequence 9  
AR000031 Sequence  
AR169142 Sequence  
AX036248 Sequence  
X52066 Clostridium  
M30196 C.botulinum  
A49989 Sequence 6  
A58945 Sequence 5  
L35496 Clostridium  
X81714 C.botulinum  
X54254 Clostridium

## ALIGNMENTS

RESULT 1  
CLOBOTB 4041 bp DNA linear BCT 26-APR-1993  
LOCUS Clostridium botulinum neurotoxin type B (botB) gene, complete cds.  
DEFINITION M81186  
ACCESSION M81186.1 GI:144734  
VERSION botB gene; neurotoxin type B.  
KEYWORDS Clostridium botulinum DNA.  
SOURCE Clostridium botulinum  
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
Clostridium.  
REFERENCE 1 (bases 1 to 4041)  
AUTHORS Whelan, S.M., Elmore, M.J., Bodsworth, N.J., Brehm, J.K., Atkinson, T.  
and Minton, N.P.  
TITLE Complete nucleotide sequence of the Clostridium botulinum gene  
encoding the type B neurotoxin  
Unpublished (1991)  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..4041  
/organism="Clostridium botulinum"  
/db\_xref="taxon:1491"  
57..3932  
gene /gene="botB"

CDS

```

57..3932
/ gene="dotB"
/ function="vertebrate neurotoxin"
/ codon_start=1
/ translat=1
/ product="neurotoxin type B"
/ protein_id="AA23211.1"
/ db_xref="GI:144735"
/ translation="MPVTINNENNDIDNNNIMMEPPFARGTRYKAKITDRIM
IIPRYEYKRPEDPKSSGIFKNDVEYDIPDLNTNDKKNLEIOTMILNPKIKK
PIGKLEMLINGIPYIGDRRPLEEENPTIASVTKLISNPEVRKRIKILNIT
PGRPVLNENETIDIGLONHFAHREGFGIMOKFCEYEVSVNNQENKASIFNRR
GYFSDPALIMHLEIHLHGLYKVDLIPNEKEFPVQSDTALQAEELTFEGD
EGKYSIDVESFDLYKSLMGPFTETIAEYKTKTASVSDSLPVKIKNLNDEY
TIEGFINSDKMEKEYRQONKAIKQAEIEIKELAVKIQMCKSVKRGICIDVD
NEDLFIADKNSFSDLSKNERLEYNTQSYIENDPIBELIDIDLSKIDLEPNT
ESITDENVDPVEKOPAKIETDENTFOYLVSOTPLDIRIDISLTSFDLLES
NKYISFSDYIKTANKVVEAGLFAWVKQIYNDPELIPVGAFLIESYIDNKKIIT
GLALNGNETAGNFENAEIAGASILEPIBELIPVGAFLIESYIDNKKIIT
DUALTRNKESDMYGLIYAOMLSTYNTOPVTEKGMVAKYQALDEIIEKYRNI
YSEKESNINIDPNDNSLNGINQALIDININFGSCVSLMKMKIDLEIIEKYRNI
DNITKKNLNTIDENKLYIGSAEYKSKYKLYKTIMPELSTYNTDITLLENKLY
NSEIILNLTIDENKLYIGSAEYKSKYKLYKTIMPELSTYNTDITLLENKLY
OMITNSVFLDFSVFIRIPKRYNDGIONTHNYTTIINCKNNSGWSISIRNRI
WTLIDINGKTSVFEYINREDISEYINRMFVDTNNLMNAKLYINKLESNTIDK
IREVLANGEIIFKLDIDIDROPTFMKYSFLENELSSNIEERYKIQSYSEYLDIP
GNPLATNKEYFMNAGNKNSTIKLKSVPSEILIRSKYINONSKYINRDYIGKPT
IRKNSOSINDIYRKEDYIYLDFFNINQEMRVYTYKFKKEEKLPLAISDSDE
YNTIOIKEDOPYSCLLFKDESDNEIDELGIIHFEISGIVFEYKDYFCISKW
YLKEYRRPYMKLKGCMNFIKPDGWE"
BASE COUNT      1679 a      383 c      645 g      1334 t
ORIGIN

```

Query Match Best Local Similarity 72.4%; Score 742.8; DB 1; Length 4041;  
Matches 963; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```

QY 10 atggccaacaatacaatcgcgaatcctcgacaacatcaccctcgacccgctgtacaa 69
Db 2610 ATGTTTAAATTAATTAATAGCGAAATTTTAAATTAATTAATTAATTAATTAAG 2669
QY 70 gacacaacatcgtacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 129
Db 2670 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2729
QY 130 gaactgaatgaagaacacggttcaactcctcgcctacactcactcactcactcgt 189
Db 2730 GAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2789
QY 190 accaagaatcagaacatcctcctcactcctcgtatcctcgtgactcctcgtcctcgt 249
Db 2790 ACTCAAAATCAGAAATATCTATTTAATAGTGTCTCGATTTCGATTTCGATTTCG 2849
QY 250 attcgtatccggaatcacaagaacggtatccagaatcactcacaatgaataaac 309
Db 2850 ATTAGAAATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 310
QY 310 atccatcactgaatgaagaataactcgtgttggaagatcctcactcgcggtacccgtatc 369
Db 2910 ATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 370
QY 370 atctggaactcgtacgtacgaacggtgaagaacaaatcgtatcctcgtacgaatac 429
Db 2970 ATATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 430
QY 430 cgtgaagaacatcctcgtacatcaatcgcgtgtcgtcgttaccatcaccaataac 489
Db 3030 AGAAGAAATATATCAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 490
QY 490 aacaatgctaataatcactcaacacggtgaactgaatcactcagaacacgaacatc 549

```

```

Db 3090 AATAACGCTAAATTTAATTAATGTAAGCTAGAACATCAATACAGATATTAAAGATAT 3149
QY 550 cgtgaagatcgcgaacggtgaagaatcactcctcaactcgaacggtgacatcgtatc 609
Db 3150 AGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 610
QY 610 cagtcacatcgtatgaataatcactcctcctcaacccggaacggtcgtacgtccatc 669
Db 3210 CAATTTATTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 670
QY 670 gaagaagaaggaataacacggtcgtacgtacgtacgtacgtacgtacgtacgtacgt 729
Db 3270 GAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 730
QY 730 ctgagtacacacagaataactcgtatggtcgtacgtacgtacgtacgtacgtacgtac 789
Db 3330 TTAATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 790
QY 790 ctgaagaagaactcctcgtgtggaatcctcgtacgtacgtacgtacgtacgtacgtac 849
Db 3390 CTAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 850
QY 850 aaatcatcaactcaccgacgtacacgtacacgtacacgtacacgtacacgtacacgtac 909
Db 3450 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 910
QY 910 aactcgtacatcactgaatgaacgtacgtacgtacgtacgtacgtacgtacgtacgtac 969
Db 3510 AATTCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 970
QY 970 ttaacactgaatcgaacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1029
Db 3570 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1030
QY 1030 aagcttctcgtcgtcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1089
Db 3630 AAATTTGTTTTTAAAGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1090
QY 1090 gaatacgaacgaacgacgacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1149
Db 3690 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1150
QY 1150 actgaagaacgtcgtcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1209
Db 3750 ACTGATGAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
QY 1210 gaataaagaactcctcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1269
Db 3810 GAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1270
QY 1270 tacaacgtgaacgtcgtcgtacgtacgtacgtacgtacgtacgtacgtacgtacgtacgtac 1329
Db 3870 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1330
QY 1330 tagtaagaat 1339
Db 3930 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1340

```

## RESULT 2

AF295926

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

1 (bases 1 to 3876)  
Klima,N., Ferreira,J.L. and Baumstark,B.R.  
Characterization of six type A strains of Clostridium botulinum

AF295926 3876 bp DNA linear BCT 02-SEP-2001  
Clostridium botulinum neurotoxin type B gene, complete cds.  
AF295926.1 GI:15419707

Clostridium botulinum.  
Clostridium botulinum.  
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
Clostridium.

gb\_pat:AX088262 + 692.00 761.78 4.6e-34 4017 ! AX088262 Sequence 10 from  
 gb\_ba:AB037704 + 683.00 752.20 1.6e-33 3835 ! AB037704 Clostridium butyr  
 gb\_ba:AB037705 + 683.00 752.20 1.6e-33 3835 ! AB037705 Clostridium butyr  
 gb\_ba:AB037706 + 683.00 752.20 1.6e-33 3835 ! AB037706 Clostridium butyr  
 gb\_ba:AB037707 + 683.00 752.20 1.6e-33 3835 ! AB037707 Clostridium butyr

seq\_name: gb\_ba:CLOBOTB

seq\_documentation\_block:  
 LOCUS CLOBOTB 4041 bp DNA linear BCT 26-APR-1993  
 DEFINITION Clostridium botulinum neurotoxin type B (botB) gene, complete cds.  
 ACCESSION M81186  
 VERSION M81186.1 GI:144734  
 KEYWORDS botB gene; neurotoxin type B.  
 SOURCE Clostridium botulinum DNA.  
 ORGANISM Clostridium botulinum  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 Clostridium.

REFERENCE  
 AUTHORS 1 (bases 1 to 4041)  
 Whelan, S.M., Elmore, M.J., Bodsworth, N.J., Brehm, J.K., Atkinson, T.  
 and Minton, N.P.

TITLE Complete nucleotide sequence of the Clostridium botulinum gene  
 encoding the type B neurotoxin  
 Unpublished (1991)

JOURNAL  
 FEATURES Location/Qualifiers  
 source 1..4041  
 /organism="Clostridium botulinum"  
 /db\_xref="taxon:1491"  
 gene 57..3932  
 /gene="botB"  
 CDS 57..3932  
 /gene="botB"  
 /function="vertebrate neurotoxin"  
 /codon\_start=1  
 /transl\_table=11  
 /product="neurotoxin type B"  
 /protein\_id="AAA23211.1"  
 /db\_xref="GI:144735"  
 /translation="MPVTINNPNYNDPIDNNNIIMPEPPFARGTGRIYKAFKITDRIW  
 IIPERYTFGYKPEDFNKSSGIFNRDVCYYDDPDYLNNDKKNIFLOTMIKLFNRIKSK  
 PLGKLLLEMIINGIPYLGDRRVPLEEFNTNIASTVTKLISNPGEVERKKGIFANLII  
 FGPGPVLNENETIDIGIQNHFAFREGFGGIMQMKFCPEYVSFNNVQENKQASIFNRR  
 GYFSDPALILMHELHVLHGLYGIKVDLLPVPNEKFFMQSTDAIQAEELYTFGGQD  
 PSIITPSTDKSIYDKVQLNFRGIVDRNLKVLVCISDPNININIKKFKDKYKFVEDS  
 EGKYSIDVESFDKLYKSLMFCFTETNIAENYKIKTRASYFSDSLPPVKIRNLLDNEY  
 TIEEGFNISDKDMEKEYRGQNKAINQAYEEISKEHLAVYKIQMCKSVKAPGICIDVD  
 NEDLFFIADKNSFDDLSKNERIEYNTQSNYIENDFPINELIOLDLISKIELPSENT  
 ES LTDNFVDPVVEKQPAIKKIFTDENTIFQYLYSQTFFPLDIRDISLTSSFDALLFS  
 NKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVNDFVIEANKSNTMDKIADISLIVPYI  
 GLALNVGNETAKGNFENAFIAGASILLEFIPELLIPVGAFLLESYIDNKKIKTI  
 DNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMYKALNYQAQALEEIKYRYNI  
 YSEKEKSNINIDFNDINSKLNENQAINQAINNINNFINGCSVSLMKMPLAVEKLLDF  
 DNTLKKNLLNYIDENKLYLIGSAEYKSVNKYKLTIMPFDLSIYNTDILIEMFNKY  
 NSEILNIIILNRYKDNLDLIDSGYAKVEYDGVDELNDKNQFKLTSSANSKIRVTQN  
 QNIIFNSVFLDFSVSEFWIRPKYKNDGIQNYIHNEYTIINCMKNNSGWKISIRGNRII  
 WTLIDINGKTKSVFFEYNIREDISYINRWFFVTITNNLNNAKIYINGKLESNTDIKD  
 IREVIANGEIIFKLDGDIRDTQFIWMKYFSIFNELSQSNIEERYKIQSYSEYKDFW  
 GNPLMYNKEYYMFNAGNKNYSYIKLKKDSPVGEILTRSKYNQNSKYINRDLYIGKFI  
 IRRKNSQSINDDIVRKEDYIYLDFFNLNQEWRVYTYKFKKEEELFLAPISDSDEF  
 YNTIQIKEYDEQPTYSCQLLFKKDEESTDEIGLIGHRFYESGIVFEEYKDYFCISKW  
 YLKEVKKPKYNLKLGCNWFIPKDEGWTE"

BASE COUNT 1679 a 383 c 645 g 1334 t  
 ORIGIN

alignment\_scores:  
 Quality: 2340.00 Length: 440  
 Ratio: 5.330 Gaps: 0  
 Percent Similarity: 99.773 Percent Identity: 99.545

alignment\_block:  
 US-09-910-186A-8 x CLOBOTB

Align seg 1/1 to: CLOBOTB from: 1 to: 4041

```

1 MetAlaAsnLysTyrAsnSerGluLeuAsnAlaLeuAsnLeu 17
2610 AGCTTTAAATAATATAGCGAAATTTAAATATATCTTAAATTT 2659
17 uAGTYrLysAspAsnLeuLeuAsnSerGlyTyrGlyAlaLysV 34
2660 AAGATATAAGCATTAATTAATTAATTAATTAATTAATTAATTA 2709
34 aLgLuAlaTyrAspGlyValGluLeuAsnAspLysAsnGlnPheLys 50
2710 TAGAGGATATGATGAGTCGACCTTAATGATTAATTAATTAATTA 2759
51 ThrSerSerAlaAsnSerLysIleArgValThrGlnAsnGlnAla 67
2760 ACTAGTTCACCAATAGTAAGATTAGAGTCACCAATATCAGATATCAT 2809
67 ePheAsnSerValPheLeuAspPheSerValSerPheTrpIleArgIleP 84
2810 ATTTAATAGTGTCTTCTTATTTAGCTTACGTTTGGATAGAATATAC 2859
84 rGlyTyrLysAsnAspGlyIleGlnAsnTyrIleHisAsnGluTyrThr 100
2860 CTAAATATAAGAAATGATGATATACAAATATATATTAATTAATTA 2909
101 IleIleAsnGlySerLysAsnAsnSerGlyTyrPheLysSerIleArg 117
2910 ATAAATTAATGATGATAAAATTAATTCGGCTGGAATAATATCTATTA 2959
117 yAsnArgIleIleTrpPheLeuIleAspIleAsnGlyLysThrLysSer 134
2960 TAAATAGCATTAATGATGATTAATTAATTAATTAATTAATTAATTA 3009
134 aPhePheGluTyrAsnIleArgGluAspIleSerGluTyrIleAsnArg 150
3010 TATTTTGAATATATACATTAAGAGATATATCAGATATATTAATTA 3059
151 TrpPhePheValThrIleThrAsnAsnLeuAsnAlaLysIleTyrIle 167
3060 TGGTTTTTGAACCTATTAATTAATTAATTAATTAATTAATTAATTA 3109
167 eAsnGlyLysLeuGlnSerAsnThrAspIleLysAspIleArgGluVal 184
3110 TAAATGTAAGCTAGAAATCAATATCAATATTAATTAATTAATTAAT 3159
184 leaIleAsnGlyLysIleIlePheLysLeuAspGlyAspIleAspArg 200
3160 TTGCTAATGCTGAATTAATTAATTAATTAATTAATTAATTAATTA 3209
201 GlnPheIleTrpMetLysTyrPheSerIlePheAsnThrGluLeuSer 217
3210 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3259
217 nSerAsnIleGlnGluArgTyrLysIleGlnSerTyrSerGluTyrLeu 234
3260 ATCAATATATTAAGAAAGATATTAATTAATTAATTAATTAATTAAT 3309
234 yAspPheTrpGlyAsnProLeuMetLysAsnLysGluTyrTyrMetPhe 250
3310 AAGATTTTGGCGAAATCCCTTAATGATCAATTAATTAATTAATTAAT 3359
251 AsnAlaGlyAsnLysAsnSerTyrIleLysLeuLysLysAspSerPro 267
3360 AATGCGGGGAAATAATTAATTAATTAATTAATTAATTAATTAATTA 3409
267 lGlyGluIleLeuThrArgSerLysTyrAsnGlnAsnSerLysTyrIle 284
3410 AGGGAATAATTTAAACACGTAGCAATATTAATTAATTAATTAATTA 3459
284 snTyrArgAspLeuTyrIleGlyGluLysPheIleIleArgArgLysSer 300
3460 ATTATAGAGATTATATATGAGAAATAATTAATTAATTAATTAATTA 3509

```

```

301 AsnSerGlnSerIleAsnAspAspIleValArgLysGluAspTyrIleTyr 317
3510 AATTCCAATCTATTAATATCATATATAGTTAGAAAAAGAAATTAATTA 3559
317 rLeuAspPhePheAsnLeuAsnGlnGluTyrPheValTyrThrTyrLysT 334
3560 TCTAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3609
334 yPheLysLysGluGluGluLysLeuPheLeuAlaProIleSerAspSer 350
3610 ATTTTAAGAAAGAGAAATAATTTGTTTACCTCTTAAAGTGAATTC 3659
351 AspGluLeuTyrAsnThrIleGlnIleLysGluTyrAspGluGlnPro 367
3660 GATGAGTTTACAAATCTATTAATTAATTAATTAATTAATTAATTAAT 3709
367 rTyrSerCysGlnLeuLeuPheLysLysAspGluGluSerThrAspGlu 384
3710 ATATAGTTCAGTGTGCTTTTAAAAAAGATGAAGAAAGTACGATGAGA 3759
384 leGlyLeuIleGlyIleHisArgPheTyrGluSerGlyIleValPheGlu 400
3760 TAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3809
401 GluTyrLysAspTyrPheCysIleSerLysTrpTyrLeuLysGluVal 417
3810 GAGTATTAAGATTAATTTTGTATTAAGTAATGCTTAATAAGAGGTAA 3859
417 sArgLysProTyrAsnLeuLysLeuGlyCysAsnTrpGlnPheIlePro 434
3860 AAGGAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3909
434 yAspGluGlyTrpThrGlu 440
3910 AAGATGAAGGCTGAGCTGAA 3929

```

seq\_name: gb\_ba:AF295926

seq\_documentation\_block: 3876 bp DNA linear BCT 02-SEP-2001

LOCUS AF295926 Clostridium botulinum neurotoxin type B gene, complete cds.

DEFINITION AF295926

ACCESSION AF295926

VERSION AF295926.1 GI:15419707

KEYWORDS

SOURCE

ORGANISM Clostridium botulinum.

REFERENCE 1 (bases 1 to 3876) Clostridium botulinum Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.

AUTHORS Kirma, N., Ferreira, J.L. and Baumstark, B.R.

TITLE Characterization of six type A strains of Clostridium botulinum that contain type B toxin gene sequences

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3876) Kirma, N., Ferreira, J.L. and Baumstark, B.R.

AUTHORS Kirma, N., Ferreira, J.L. and Baumstark, B.R.

TITLE Direct Submission

JOURNAL Submitted (14-AUG-2000) Department of Biology, Georgia State University, P.O. Box 4010, Atlanta, GA 30302-4010, USA

FEATURES

source

1..3876 Location/Qualifiers

1..3876 /organism="Clostridium botulinum"

/isolate="1436"

/db\_xref="taxon:1491"

/country="USA: Utah"

/note="Isolated from stool sample in 1977"

type: "Ab"

1..3876 /codon\_start=1

/transl\_table=1

/product="neurotoxin type B"

/protein\_id="AAK97132.1"

/db\_xref="GI:15419708"